ABSTRACT: Spatial memory in mammals, including humans, appears highly sexually dimorphic. The present investigation sought to examine if spatial learning and spatial memory in humans is also linked to sexual orientation. This was achieved by using virtual reality versions of two classic paradigms developed in animal models of hippocampal functioning, the Morris Water Maze (MWM) and Radial Arm Maze (RAM). Here, we show that in contrast to heterosexual men, and in congruence with heterosexual women, homosexual men displayed significantly greater search latencies (spatial learning) during a virtual Morris Water Maze. During a virtual 8-arm Radial Arm Maze, heterosexual males had significantly shorter search latency than homosexual females, and did not differ from homosexual males. Statistical modeling revealed that variations in neurodevelopmental markers previously associated with human sexual orientation (2nd to 4th finger length ratios and older fraternal siblings) differentially predicted MWM probe trial performance and RAM search latencies only. These data may limit the number of possible neurodevelopmental pathways responsible for sexual variation in components of spatial learning and memory.

KEY WORDS: sex differences; homosexuality; Morris Water Maze; Radial Arm Maze; hippocampus

INTRODUCTION

Sexual dimorphisms in spatial memory are amongst the most robust individual differences reported in humans and shows large effect sizes (Voyer et al., 1995; Astur et al., 1998). Males outperform females, on average, on tests of place learning and navigation whereas females outperform males, on average, on tests of spatial location memory where object-to-place binding is salient (Kimura, 1999). Most behavioral investigations of male-favoring place learning rely on route learning-type tasks which mainly elicit egocentric frames of spatial reference (viewer-dependent), and these demonstrate a variable male advantage (in paper-and-pencil tests: Galea and Kimura, 1993; Dabbs et al., 1998; Choi and Silverman, 1996; tasks using photographs: Holding and Holding, 1989; real-life terrain level routes: Saucier et al., 2002; Silverman et al., 2000; and computer simulations using “mazes”: Moffat et al., 1998).

However, notable exceptions are 3D virtual reality human analogues of classical paradigms from animal models of hippocampal functioning, namely the Morris Water Maze (MWM) and Radial Arm Maze (RAM) tasks. These can be experimentally manipulated and tap mainly allocentric frames of reference (especially MWM tasks; O’Keefe and Nadel, 1978; Morris and Mayes, 2004). Studies report large and robust sex differences on two MWM outcome measures with males (humans and rodent) demonstrating (i) shorter latencies (better spatial learning) to locate hidden platforms and (ii) greater time spent in the training quadrant (greater dwell time) of the tasks compared with females (Astur et al., 1998, 2004; Driscoll et al., 2005; Jonasson, 2005). Human males have also been found to show shorter latencies than females on the RAM in one study but not in another (Iaria et al., 2003; cf. Astur et al., 2004).

In terms of neural correlates, several human lesion studies report spatial learning and memory impairments following hippocampal damage on MWM and RAM analogues (computer-simulated and nonsimulated versions, e.g., Abrahams et al., 1997, 1999; Astur et al., 2002; see also Kessels et al., 2001 for a review). Functional neuroimaging studies in healthy humans also point to bilateral hippocampal involvement during MWM spatial learning (e.g., Parlow et al., 2004) although both functional and volumetric data regarding sexual dimorphism are inconclusive (Filipek et al., 1994; Gron et al., 2000 cf. Good et al., 2001; Blanch et al., 2004; Janzen and Van Turennot, 2004).

Sex differences in human spatial cognition are further confused by several independent reports of within-sex variation attributable to sexual orientation. Studies have consistently demonstrated that healthy adult heterosexual males out-perform homosexual males on male-favoring tests such as mental rotation, the water level test, and judgement of line orientation (Glade et al., 1990; McCormick and Witelson, 1991; Wegesin, 1998; Neave et al., 1999; Rahman and Wilson, 2003). Homosexual males also performed better than heterosexual males in one study of a female-favoring spatial location memory test, suggesting some domain-specific dissociation of spatial processes (Rahman et al., 2003b). Moreover Rahman et al. (2005) reported that homosexual males used a significantly greater landmark-based navigation strategy (female-favoring) than did heterosexual males. The
effect sizes for the heterosexual and homosexual male contrasts are large suggesting they are nontrivial group differences. Homosexual females do not differ from heterosexual females in most neurocognitive functions except on frontally-mediated verbal fluency tests, where the former group show male-typical performance (Rahman et al., 2003a).

The neurodevelopment of both sex and sexual orientation-related differences in spatial cognition has been proposed to depend on prenatal androgen levels (Collaer and Hines, 1995; Jonasson, 2005). Homosexuals are predicted to show sex-atypical neurocognitive patterns, consistent with the atypical shift in their sexual partner preference, under the actions of these prenatal androgens (Ellis and Ames, 1987; Rahman, 2005). Thus far, the pattern of cross-sex shifts in spatial cognition among homosexual males (and in verbal fluency among homosexual females) is supportive. As invasive measurements of prenatal sex steroids are impractical for comparative studies in adults, current work utilizes certain “proxy” somatic markers. Manning (2002) summarized the evidence that androgens stimulates prenatal growth of the 4th finger while prenatal estrogen stimulates the growth of the 2nd finger—a low 2nd to 4th finger length ratio (or 2D:4D) being indicative of greater exposure to androgens prenatally. In support, 2D:4D is fixed as early as 2 yr, and the ratio is higher in females than in males (Manning, 2002); males and females with congenital adrenal hyperplasia (who are exposed to high levels of adrenal androgens in utero) have low values of 2D:4D compared with controls (Brown et al., 2002); 2D:4D is linked to sequence variation in the androgen receptor gene (Manning et al., 2003); and at 2 yr of age 2D:4D is negatively associated with the amniotic testosterone (cf. estrogen) levels sampled during gestation (Lutchmaya et al., 2004). Moreover, 2D:4D among human females is correlated with volumetric asymmetry of the hippocampus and improved place learning (Csatho et al., 2003; Kallai et al., 2005). Homosexual males and females also show lower (i.e., masculinized) 2D:4D ratios compared with heterosexuals, and although the effect for males appears present in British studies only (for review see McFadden et al., 2005; Rahman, 2005), these data indicate that homosexuals of both sexes may be exposed to elevated prenatal androgens. One final neurodevelopmental consideration is extensive evidence for a “fraternal birth order” (FBO) effect in which homosexual males evidence later birth order in relation to their brothers only (Blanchard, 2004). Blanchard (2004) reviews evidence proposing that this effect may be due to progressive maternal immunization of male-linked antigens with each male fetus, shifting neural sexual differentiation of successive male fetuses in a feminized direction. Furthermore, Blanchard suggests that FBO may correlate with spatial cognitive ability among homosexual males or males in general.

The present study sought to examine, for the first time, sexual orientation-related performance differences in known spatial learning and memory outcome measures on computer simulated versions of the MWM and 8-arm RAM. On the basis of the extant literature, we predicted that heterosexual females and homosexual males would show significantly greater latencies to find a hidden platform and spend less dwell time in the training quadran of the MWM compared with heterosexual males. Given the inconsistent findings regarding sex differences on the RAM, no group differences were predicted here. Also, no differences were predicted between homosexual and heterosexual females on any task. This study also investigated the novel hypothesis that neurodevelopmental markers previously associated with human sexual orientation (2D:4D and FBO) may be associated with performance variations, thus narrowing the number of putative neurodevelopmental explanations for sexual variation in allocentric spatial learning and memory.

**MATERIALS AND METHOD**

**Participants**

One hundred and forty participants (aged 19–45) took part (35 heterosexual males, 35 heterosexual females, 35 homosexual males, 35 homosexual women). All potential participants for the study were screened before taking part to exclude any history of psychiatric or neurological morbidity, head injury, and use of psychoactive medication or drug abuse (one heterosexual male, one heterosexual female, and two homosexual males). Data from participants reporting motion sickness during the virtual tasks was also excluded (three heterosexual females and three homosexual males) and further participants were recruited to replace those lost. Heterosexuals and homosexuals were recruited via advertisements (in community newspapers for heterosexual participants and gay/lesbian press for homosexual participants) followed by snowball sampling whereby the initial wave responding to the adverts were asked to solicit others to take part (an established method for increasing size of sexual minority samples). Sexual orientation was assessed using self-identification (“gay/lesbian,” “heterosexual/straight” or “bisexual”) and a single-item question about sexual attractions and fantasies on a 7-point scale (ranging from 0 = “exclusively heterosexual” to 6 = “exclusively homosexual”). Only participants who responded either 0 and 1 (heterosexual) and 5 or 6 (homosexual), and checked either “gay/lesbian” or “heterosexual/straight” on self-identification took part (bisexual respondents were excluded). Written informed consent was obtained from all participants and the UEL Ethics Committee approved all procedures.

**Measures and Procedure**

Four neurocognitive tasks were used (administered randomly across subjects): the Wechsler test of adult reading, mental rotation, a virtual MWM, and a virtual 8-arm RAM. Two neurodevelopmental predictors, right- and left-hand 2D:4D, and sibling sex composition, were also used.

**Wechsler test of adult reading (WTAR)**

An estimate of general cognitive ability was provided by the WTAR (WAIS III full-scale IQ [FSIQ] scores).
Mental rotation

This 10 min, 20-item test (Vandenberg and Kuse’s, 1978 adaptation from Shepard and Metzler, 1971) required viewing and matching target stimuli (a 2D representation of a 3D cuboid) with four test stimuli. Each item had two correct test stimuli and two incorrect foils. For each item, participants received 2 points if they checked both correct responses and 1 point if they checked one correct response. The maximum score was 40. This basic spatial ability test shows a male advantage and correlates with MWM probe trial performance (see Astur et al., 2004).

Virtual Morris Water Maze (MWM)

For full published protocols see Astur et al. (1998, 2004). In brief, participants were simply instructed that they would find themselves in a virtual pool which they had to escape from as quickly as possible by swimming (using keyboard arrow keys) to a hidden platform using extra-pool cues (icons) for guidance. The PC monitor displayed a 60° field of view from the first-person perspective, approximately at the same level as the eye. The icons were three wall pictures, a large television, an up-right lamp, three doorways, a sofa, a telephone upon a stand, and a bookcase. The arena walls and ceiling were not textured but the rim of the pool was. During a “hidden platform training” phase participants started from four different cardinal locations, four times (randomized) over 16 trials (74 s each) with a 5 s ITI. If a participant swam over the hidden platform a pleasant tone sounded indicating success. At this point, the participant could look around for 5 s from their position upon the platform after which the trial was terminated, the maze reset and participants were required to find the platform again. Subsequently, a probe trial was given in which the platform was removed and free search permitted to find the platform for 30 s, after which the trial was terminated (no indication was provided that this trial was different from the previous phase). Finally, a “visible platform” phase required participants to swim to a raised platform, starting from four different cardinal points for four trials.

Virtual 8-arm RAM

In brief (see Astur et al., 2004), this task required participants to traverse 8 “arms” extending from a circular junction, four of which contained rewards and four which did not. The room had several icons including a plant in a pot, three wall pictures, two up-right room lamps, a bureau (table, chair and computer), a bookcase, a table with a lamp upon it, a sofa, a doorway, and two Corinthian style pillars. The walls and ceiling in this room were textured. Participants were instructed that they had to retrieve all four rewards as quickly as possible by “walking” up the arms (successful retrieval of all four indicated by a pleasant tone and the message “reward” when the participant walked over the reward). They were also instructed to avoid reentering an arm in which they had already retrieved a reward. Once participants had found all the rewards a message displaying “congratulations” was presented and participants were transported back to the circular junction to begin the next trial after an ITI of 5 s (over a total of 10 trials). An elapse of 3 min without reward retrieval led to trial termination. The sequence of reward arms varied among participants. A reference memory error was scored if participants entered an unrewarded arm (trial-independent information remained constant throughout testing; Olton and Pappas, 1979), and working memory errors were scored if they entered a previously visited arm during that trial, irrespective of reward presence (information about which rewards had been previously retrieved varied from trial to trial; Olton and Pappas, 1979).

Neurodevelopmental predictor variables

Briefly (see Manning, 2002 for full protocol), the lengths of second and fourth digits on each hand were measured twice, using digital callipers (measuring to 0.01 mm), on the ventral surface of the hand, from the basal crease of the digit to the tip (the most proximal crease is measured where several creases exist). Our measurements showed high repeatability (Cronbach’s α ranging from 0.97 to 0.99), consistent with those shown previously for finger digit measures (Manning, 2002). The measurements were averaged, and the ratios calculated by dividing 2nd digit length by 4th digit length (= 2D:4D) for right and left hands separately. Participants also listed the numbers of older brothers, younger brothers, older sisters, and younger sisters that their mothers had carried. Only biological siblings were to be listed, including those carried but not brought to term because of obstetric complications.

Statistical Analysis

Except for categorical variables (analyzed using χ²), measures were analyzed using factorial (sex by sexual orientation) ANOVA. For virtual spatial memory measures, factorial ANCOVA with covariates and repeated measures, where appropriate (see Results), were used. Significant interactions were unpacked using three t-tests with a Bonferroni correction (P < 0.01). All other α’s were at 0.05. The effect sizes for these contrasts is also reported according to Cohen’s d, where d = 0.2 is regarded as a small effect, d = 0.5 a medium effect, and d ≥ 0.8 a large effect (Cohen, 1988). To examine the contribution of putative predictors to sexual variation in spatial memory, multiple regressions were conducted between significant performance measures and the predictor variables. A dummy coded variable (“group”), comprising homosexual males, heterosexual females, and homosexual females, with heterosexual males as the reference group, was forced entered into the first block of the model permitting between-group differences to be attributed to “group.” Predictor variables (age, years in education, estimated FSIQ, mental rotation score, right- and left-hand 2D:4D, and number of older brothers) were then entered stepwise into a second block (reported in the text are effect sizes r², adjusted r², and r² change for each step).
TABLE 1.

Means (SD) for Participant Characteristics, Estimated FSIQ, Mental Rotation, and Neurodevelopmental Predictors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Heterosexual males</th>
<th>Heterosexual females</th>
<th>Homosexual males</th>
<th>Homosexual females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>29.37 (6.95)</td>
<td>27.34 (6.79)</td>
<td>32.63 (8.34)</td>
<td>29.03 (6.75)</td>
</tr>
<tr>
<td>Years in education</td>
<td>15.85 (2.70)</td>
<td>16.60 (2.41)</td>
<td>13.87 (3.21)</td>
<td>14.08 (4.08)</td>
</tr>
<tr>
<td>Handedness scores&lt;sup&gt;a&lt;/sup&gt;</td>
<td>60.53 (57.51)</td>
<td>73.99 (37.61)</td>
<td>67.27 (54.40)</td>
<td>54.68 (45.18)</td>
</tr>
<tr>
<td>Perceived stress&lt;sup&gt;b&lt;/sup&gt; after WTAR</td>
<td>3.71 (2.48)</td>
<td>3.48 (3.21)</td>
<td>2.42 (2.07)</td>
<td>4.34 (2.87)</td>
</tr>
<tr>
<td>Perceived stress after Mental rotation</td>
<td>4.26 (1.99)</td>
<td>4.67 (1.76)</td>
<td>4.43 (2.55)</td>
<td>4.63 (2.69)</td>
</tr>
<tr>
<td>Perceived stress after virtual MWM</td>
<td>2.40 (1.71)</td>
<td>3.00 (2.15)</td>
<td>2.94 (2.20)</td>
<td>2.89 (2.37)</td>
</tr>
<tr>
<td>Perceived stress after virtual 8-arm RAM</td>
<td>2.23 (2.60)</td>
<td>2.80 (2.65)</td>
<td>2.40 (2.26)</td>
<td>2.23 (2.17)</td>
</tr>
<tr>
<td>Computer game experience&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.97 (3.01)</td>
<td>1.40 (1.59)</td>
<td>1.97 (2.43)</td>
<td>2.54 (2.46)</td>
</tr>
<tr>
<td>Estimated FSIQ scores&lt;sup&gt;d&lt;/sup&gt;</td>
<td>106.72 (7.74)</td>
<td>110.36 (6.33)</td>
<td>108.56 (5.75)</td>
<td>107.38 (6.95)</td>
</tr>
<tr>
<td>Mental rotation scores</td>
<td>28.06 (5.65)</td>
<td>25.38 (8.09)</td>
<td>25.59 (7.18)</td>
<td>23.66 (6.27)</td>
</tr>
<tr>
<td>Right-hand 2D:4D</td>
<td>0.97 (0.03)</td>
<td>0.98 (0.03)</td>
<td>0.97 (0.03)</td>
<td>0.97 (0.02)</td>
</tr>
<tr>
<td>Left-hand 2D:4D</td>
<td>0.98 (0.02)</td>
<td>0.99 (0.03)</td>
<td>0.97 (0.03)</td>
<td>0.98 (0.2)</td>
</tr>
<tr>
<td>Number of older brothers</td>
<td>0.34 (0.53)</td>
<td>0.26 (0.44)</td>
<td>0.63 (0.84)</td>
<td>0.43 (0.60)</td>
</tr>
<tr>
<td>Number of younger brothers</td>
<td>0.37 (0.49)</td>
<td>0.37 (0.54)</td>
<td>0.60 (0.91)</td>
<td>0.29 (0.45)</td>
</tr>
<tr>
<td>Number of older sisters</td>
<td>0.46 (0.56)</td>
<td>0.51 (0.50)</td>
<td>0.60 (1.03)</td>
<td>0.57 (1.19)</td>
</tr>
<tr>
<td>Number of younger sisters</td>
<td>0.29 (0.51)</td>
<td>0.26 (0.50)</td>
<td>0.40 (0.55)</td>
<td>0.29 (0.57)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Measured via the Edinburgh handedness inventory (EHI).
<sup>b</sup>Measured from 0 (not at all stressed) to 10 (extremely stressed).
<sup>c</sup>Measured from 0 (never) to 10 (frequently).
<sup>d</sup>Estimated WAIS III full-scale IQ from WTAR standard scores.

**RESULTS**

**Participant Characteristics**

There were no significant group differences in handedness scores (evaluated using the Edinburgh Handedness Inventory; Oldfield, 1971), perceived stress levels during each task, and computer game playing experience (all $P_{s} < 0.05$. See Table 1). There were no differences in ethnicity (“white” vs. “nonwhite”: $\chi^2(3) = 3.182, P = 0.364$). The female groups did not differ in oral contraceptive use or menstrual cycle phase at testing (“menstrual phase” = 2–5 days from the onset of last menstruation; “mid-luteal phase” = 5–10 days counting backwards from the onset of the next menstruation; “other” = all other days; all $P_{s} > 0.05$). However, males were significantly older than females $[F(1, 139) = 5.284, P = 0.023$, a significant block by sex by sexual orientation interaction $[F(3, 119) = 2.931, P = 0.031$, but no other significant within-subjects interactions $[all P_{s} > 0.05]$. There was a significant between-subjects effect of sex $[F(1, 121) = 16.461, P < 0.001]$, males showing shorter latencies to find the hidden platform than females, no significant effect of sexual orientation $[F(1, 121) = 0.626, P = 0.405]$, and a significant sex by sexual orientation interaction $[F(1, 121) = 4.947, P = 0.028]$ (Table 2 and Fig. 1.).

There was a significant interaction for estimated FSIQ $[F(1, 131) = 4.204, P = 0.042]$, however, post-hoc comparisons were not significant at the conservative $\alpha$ level, all $P_{s} > 0.05$. See Table 1).

**Estimated FSIQ Scores**

There was a significant interaction for estimated FSIQ $[F(1, 131) = 4.204, P = 0.042]$, however, post-hoc comparisons were not significant at the conservative $\alpha$ level, all $P_{s} > 0.05$. See Table 1).

**Mental Rotation**

A main effect of sex, revealing that males tended to score higher than females, just failed to reach significance $[F(1, 135) = 3.870, P = 0.051, d = 0.34]$. There were no other significant effects $[all P_{s} > 0.05$. See Table 1].

**Virtual MWM**

Latencies across the trials were collapsed into four blocks. Factorial ANCOVA (age, years in education, and estimated FSIQ as covariates) with block as a repeated measure revealed no significant effect of block $[Wilk’s F(3, 119) = 0.624, P = 0.601]$, a significant block by sex by sexual orientation interaction $[Wilk’s F(3, 119) = 3.052, P = 0.031]$, but no other significant within-subjects interactions $[all P_{s} > 0.05]$. There was a significant between-subjects effect of sex $[F(1, 121) = 16.461, P < 0.001]$, males showing shorter latencies to find the hidden platform than females, no significant effect of sexual orientation $[F(1, 121) = 0.602, P = 0.803]$, and a significant sex by sexual orientation interaction $[F(1, 121) = 4.947, P = 0.028]$ (Table 2 and Fig. 1.).

To unpack the interaction the blocks were averaged and subjected to post-hoc tests which revealed significantly shorter latencies in heterosexual males compared with heterosexual females $[t(68) = -3.120, P = 0.003, d = 0.72]$ and homosexual males $[t(68) = 2.931, P = 0.005, d = 0.81]$. There

_Hippocampus_ DOI 10.1002/hipo
were no significant differences between heterosexual and homosexual females \[t(68) = -0.675, \ P = 0.502, \ d = 0.05\].

There were no significant within- or between-subjects effects in distances travelled to find the hidden platform when trials were collapsed into four blocks or in average distance travelled across all trials \[all \ P_s > 0.05\]. Similarly, no terms were significant for swim latencies and distances travelled during the visible platform phase \[all \ P_s > 0.10\]. However, there was a significant main effect of sex \[but not any other term; \ P_s > 0.05\] during the probe trial for percentage latency \[\text{Wilk}'s \ F(1, 128) = 18.172, \ P < 0.001, \ d = 0.55\] and percentage distance \[\text{F}(1, 128) = 16.473, \ P < 0.001, \ d = 0.49\] spent in the training quadrant, males spending more dwell time and greater distance than females. Note here that all groups performed above chance \(0.25\) and spent the majority of their dwell time and distance in the correct quadrant (Table 2).

### Virtual 8-Arm RAM

With trial as the repeated measure, ANCOVA revealed no significant within-subjects effect of trial \[\text{Wilk}'s \ F(9, 106) = 1.179, \ P = 0.316\], a significant within-subjects trial by sex interaction \[\text{Wilk}'s \ F(9, 106) = 2.707, \ P = 0.007\], and a significant within-subjects trial by sex by sexual orientation interaction \[\text{Wilk}'s \ F(52, 576) = 2.215, \ P = 0.023\] because of a trend for homosexual males showing shorter latency than females, no significant effect of sexual orientation \[\text{F}(1, 114) = 0.006, \ P = 0.936\], and a significant sex by sexual orientation interaction \[\text{F}(1, 114) = 7.691, \ P = 0.006\] (see Table 3 for the averaged measure across the 10 trials and Fig. 2 for acquisition curves). Heterosexual males had significantly shorter latency than homosexual females \[t(68) = -2.570, \ P = 0.010, \ d = 0.61\]. There was no significant difference between the two male groups \[t(68) = 1.515, \ P = 0.134, \ d = 0.47\] and there was a nonsignificant trend for shorter latency in homosexual females compared with heterosexual females \[t(68) = -2.323, \ P = 0.023, \ d = 0.56\]. Note Figure 2 also shows that group differences tail off by the later trials (trials 8–10). There were no significant effects for distances travelled or working memory errors (logarithmic transformation was applied to these variables because of skewed distributions; all \(P_s > 0.05\)). However, there was a significant interaction for reference memory errors \[\text{F}(1, 129) = 5.064, \ P = 0.026\] because of a trend for homosexual males to make fewer errors than heterosexual females \[t(68) = -2.215, \ P = 0.030, \ d = 0.56\].

### Regression Modeling of Sexual Orientation-Related Spatial Memory Differences

There were no group differences in right-hand 2D:4D or left-hand 2D:4D \(P_s > 0.05\) (Table 1), but homosexual males showed the expected trend (nonsignificant) towards greater numbers of older brothers compared with heterosexual males \[t = 1.689, \ df = 68, \ P = 0.096, \ d = 0.41\]. Comparisons for any other class of sibling did not approach significance \(P_s > 0.10\).

For MWM hidden platform search latencies, group explained 10% \[\text{adjusted} \ r^2 = 0.078\] of the variance. The addition of age at the second step led to a significant \(r^2\) change increase by 17.4%. Mental rotation scores at the third step, and years in education at the fourth step, explained a significant 4.9% and 2.4% respectively. In the final step, estimated FSIQ explained a further 2.4% of the variance and the com-

<table>
<thead>
<tr>
<th>Variable</th>
<th>Heterosexual males</th>
<th>Heterosexual females</th>
<th>Homosexual males</th>
<th>Homosexual females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search latency block 1 (s)</td>
<td>34.59 (15.12)</td>
<td>47.05 (17.81)</td>
<td>44.76 (15.66)</td>
<td>42.88 (15.57)</td>
</tr>
<tr>
<td>Search latency block 2 (s)</td>
<td>22.32 (19.91)</td>
<td>30.31 (13.16)</td>
<td>34.44 (17.29)</td>
<td>29.76 (15.57)</td>
</tr>
<tr>
<td>Search latency block 3 (s)</td>
<td>19.09 (15.42)</td>
<td>30.12 (12.34)</td>
<td>29.21 (16.50)</td>
<td>28.49 (12.54)</td>
</tr>
<tr>
<td>Search latency block 4 (s)</td>
<td>19.28 (14.09)</td>
<td>26.89 (14.85)</td>
<td>27.35 (15.70)</td>
<td>30.53 (16.37)</td>
</tr>
<tr>
<td>Average distance (au)</td>
<td>1171.50 (488.37)</td>
<td>1364.13 (428.28)</td>
<td>1721.31 (855.37)</td>
<td>1557.62 (717.64)</td>
</tr>
<tr>
<td>Visible trials search latency (s)</td>
<td>8.68 (2.46)</td>
<td>9.46 (3.46)</td>
<td>18.83 (46.16)</td>
<td>8.78 (3.67)</td>
</tr>
<tr>
<td>Visible trials distance (au)</td>
<td>340.56 (88.33)</td>
<td>334.20 (72.45)</td>
<td>434.89 (86.33)</td>
<td>333.84 (66.32)</td>
</tr>
<tr>
<td>% latency in training quadrant (probe)</td>
<td>58.30 (18.51)</td>
<td>39.84 (28.62)</td>
<td>51.48 (20.86)</td>
<td>44.19 (24.72)</td>
</tr>
<tr>
<td>% distance in training quadrant (probe)</td>
<td>58.60 (16.38)</td>
<td>41.77 (28.26)</td>
<td>51.35 (17.97)</td>
<td>46.21 (23.51)</td>
</tr>
</tbody>
</table>

**FIGURE 1.** Acquisition curves for search latency during the virtual MWM across groups. Heterosexual males swim to the hidden platform significantly more quickly than the other groups.
bined variables explained 37.2% [adjusted $r^2 = 0.335$; $F(7, 127) = 10.137, P < 0.001]$ of the overall variance (Table 4). Rejected variables were number of older brothers, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001]$. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D. For percentage latency spent in the training quadrant, group explained 6.9% [adjusted $r^2 = 0.047$] of the variance at the first step, mental rotation explained a further 27.2% at the second step, age explained 5.2% at the third step, and number of older brothers explained 2.4% at the fourth step. A total of 41.7% [adjusted $r^2 = 0.388$; $F(6, 126) = 14.303, P < 0.001$] was explained by the combined variables at this final step (rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D).

For virtual RAM search latency, group explained 6.3% of the variance, followed by mental rotation [5.65 (2.53)] at the second step. Adding age [3.72 (0.74)] at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D. For percentage latency spent in the training quadrant, group explained 6.9% [adjusted $r^2 = 0.047$] of the variance at the first step, mental rotation explained a further 27.2% at the second step, age explained 5.2% at the third step, and number of older brothers explained 2.4% at the fourth step. A total of 41.7% [adjusted $r^2 = 0.388$; $F(6, 126) = 14.303, P < 0.001$] was explained by the combined variables at this final step (rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D).

For virtual RAM search latency, group explained 6.3% of the variance, followed by mental rotation [5.65 (2.53)] at the second step. Adding age [3.72 (0.74)] at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.

For percentage latency spent in the training quadrant, group explained 7.9% [adjusted $r^2 = 0.056$] of the variance, followed by mental rotation score [$r^2$ change = 26.3%] at the second step. Adding age at the third step led to a significant $r^2$ change increase of 4.3% and number of older brothers explained a further 2%. At this final step, 40.4% of the variance was explained by the combined variables [adjusted $r^2 = 0.374$; $F(6, 126) = 13.562, P < 0.001$]. Rejected variables were estimated FSIQ, years in education, right- and left-hand 2D:4D.
During a virtual MWM, heterosexual males demonstrated shorter search latencies compared with the other groups consistent with our predictions. There were no group differences in distances travelled and males (irrespective of sexuality) spent more of their dwell time and distance in the training quadrant during a probe trial (with all groups spending most of their time overall here). During a virtual RAM, heterosexual males were also faster compared with heterosexual females; were no different to homosexual males, while heterosexual and homosexual females did not differ significantly here at a conservative α value. Examination of the effect sizes across the contrasts confirms that these between-group differences are modest to large in their magnitude but also show that they occur for some spatial learning and memory outcome measures and not others.

Observation of the acquisition curves (Figs. 1 and 2) shows clearly that heterosexual males started faster on both tasks and maintained this advantage throughout the learning trials. This suggests that heterosexual males may be using unique search strategies from the first trial in each task compared with the other groups. Importantly, in both tasks we found group differences in terms of latency but none in terms of distances travelled. This also suggests that heterosexual men may be equally efficient to the other groups in navigation but different to homosexual males, while heterosexual and homosexual females did not differ significantly here at a conservative α value. Examination of the effect sizes across the contrasts confirms that these between-group differences are modest to large in their magnitude but also show that they occur for some spatial learning and memory outcome measures and not others.

Overall data appear to support a “mosaic” view that any sexual orientation-related differences will be component-specific rather than representing global cognitive shifts in either male- or female-typical directions (Rahman, 2005). The results also showed a human sex difference favoring males in the RAM which might be due to greater power of the present study over others using the same task but reporting null findings (e.g. Astur et al., 2004). This is consistent with a small, yet reliable male advantage on the RAM in mouse models (Jonasson, 2005). It should be pointed out that any suggestion of “mosaic” or component-specific cognitive processing implies little about the mechanisms causally involved in this manifestation (and the findings regarding neurodevelopmental markers are discussed below) and might simply reflect the unsystematic nature of sexual variation in spatial learning and memory we observed here.

The MWM task employed here has previously been shown to depend on hippocampal, rather than extra-hippocampal, integrity (Astur et al., 2002). However, the pattern of findings reported here do not directly speak to hippocampal involvement therein. We have already suggested that sexual orientation may affect navigating speed towards a target but that this does not affect the overall spatial knowledge that arises. Also, heterosexual males may be utilizing a navigation strategy that is not moderated by the hippocampus, although the hippocampus may be involved in encoding of allocentric spatial representations (Kessels et al., 2001). Moreover, it could be argued that shorter search latencies [among heterosexual males] across the tasks reflects reduced allocentric processing because of differential attentional modulation by the number of extra-maze cues. Without testing several mazes where the number of such cues is systematically varied it is impossible to examine whether search latency is correlated with the size of a cognitive map. In any case, we cannot conclude that the performance differences in search latency observed are subtle indicators of differences in hippocampal structure and/or functioning.

A secondary aim of this study was to elucidate which neurodevelopmental factors contribute to any sexual variation in spatial behavior. Regression modeling revealed that markers of prenatal androgen exposure (2D:4D; Manning, 2002) predicted variance in RAM search latency only. This is consistent with studies showing that castrated male rats display poorer RAM performance while female rats treated with estradiol benzoate or testosterone perform better than controls (Williams et al., 1990; Roof, 1993). The finding for an influence of number of older brothers on probe trial performance is intriguing. At this stage, the mechanisms underlying this relationship remain unclear although they may involve maternal immunization responses proposed for the FBO effect in male sexual orientation (Blanchard, 2004). However, no effect of sexual orientation was demonstrated for probe trial measures and the amount of variance explained by the addition of older brothers was small. Importantly, both 2D:4D and number of older brothers (traits previously associated with human sexual orientation) contributed no independent predictive power to the group differences reported for MWM search latency. This
finding is perplexing given that prenatal manipulation of androgen levels (or examination of androgen receptor mutants compared to normal animals) produces cross-sex shifts in search latencies during water maze protocols by male animals (Williams et al., 1990; Jones and Watson, 2005; Risk et al., 2005). It is possible that of the number of neurodevelopmental predictors of sexual orientation in humans, the present study examined a very narrow range of measures (note the lack of group differences reported here). However, it also likely that our two neurodevelopmental markers are, in fact, poor indices of the developmental processes they are purported to reflect. For example, the sex difference in 2D:4D is typically less than half a standard deviation in population level samples whereas sex differences in prenatal androgen levels are very large ( Manning, 2002). Similarly, while the FBO effect is purported to reflect maternal immunity processes, there are no biochemical studies of this in humans. Further work will need to take account of multiple factors.

While we did not directly investigate age-related differences in this study, age did account for a significant and large proportion of the variation in both MWM and RAM search latency scores. This is consistent with previous studies demonstrating an association between age and place learning in the form of age-related declines across a range of allocentric tasks, but also that the male advantage on search latency (during the MWM at least) persists (e.g., Driscoll et al., 2005; Moffat and Resnick, 2002). In this study, males were older but showed better performance than females, while homosexuals (as a group) were older than heterosexuals but homosexual males showed poorer performance relative to heterosexual males on the MWM. Given this complex pattern, future studies should examine interactions between age and sexual orientation in spatial behavior.

The observation that mental rotation scores predicted substantial variance in each spatial memory measure partly confirms other work, e.g., Astur et al. (2004) found a similar relationship for the MWM but not the RAM. Our data suggest that some male-typical spatial behavior in healthy humans may arise from other facets of basic spatial working memory. Importantly, even though the RAM and MWM are widely and interchangeably used measures of spatial behavior in nonhuman animals and in human lesion studies, the pattern of results reported here may indicate that mental rotation and MWM performance are in fact better measures of spatial behavior overall. Moreover, the spatial demands of the MWM may be slightly greater than those of the RAM which may contribute to a male advantage in spatial learning. However, this comment must be placed in context of the fact that no statistically significant male-advantage on mental rotation was demonstrated (and neither did we find that homosexual males performed poorer relative to heterosexual males as demonstrated in prior work e.g., Rahman and Wilson, 2003). We have no explanation for this at present other than to note the sex difference was in the expected direction (although the effect size was small) and just failed to each significance. Our sample might comprise individuals with high overall cognitive ability which may mask individual differences in this test of basic spatial ability.

In conclusion, the comparative differences reported constitute medium to large effect sizes and show nontrivial sexual orientation-related variation in certain measures of spatial learning and memory. Nonetheless, future work must now clarify the extent of these variations, the navigation strategies used, and any role for the hippocampus using neuroimaging studies with virtual protocols robustly sensitive to behavioral differences. Moreover, such protocols should consider burgeoning work pointing to allocentric spatial processing being achieved via mental transformation of viewpoints within 3D space applied to viewer-dependent (egocentric) representations (e.g., King et al., 2002; Parslow et al., 2004). This will aid in narrowing the number of underlying cognitive processes involved in putative sexual variation.

REFERENCES

SEXUAL ORIENTATION AND ALLOCENTRIC SPATIAL MEMORY


